EXACT DISTRIBUTION AND PERFORMANCE PROPERTIES OF PRE-TEST ESTIMATOR OF REGRESSION COEFFICIENT UNDER ORTHONORMAL REGRESSION MODEL

Vinod Kumar
Department of Mathematics, Ch. Bansi Lal University, Bhiwani

ABSTRACT

The article studies the exact distribution of Pre-test estimator of regression coefficient under orthonormal regression model. The performance properties of Pre-test estimator have been studied empirically. Attempt has also been made to carry out a comparative study of Pre-test estimator with ordinary least squares and restricted least squares estimators.

Keywords: Ordinary Least Squares Estimator (OLS), Restricted Least Squares Estimator (RLS), Pre-Test Estimator, Quadratic Error Loss.

AMS Mathematics Subject Classification (2010) - 62J05

1. INTRODUCTION

In regression analysis, the main objective is to extract the maximum amount of information from the data. In fact explicit determination of coefficients enhances when there is some non-sample information together with the available sample information. Such information may be available from theoretical postulations, empirical experiences or from some other sources, and are incorporated in the model in the form of restrictions on the parameters. Innumerable examples may be sighted from economics and other applied sciences where some prior information binding the parameters of the model may be available. For instance while working with Cobb-Douglas Production function, the information that a particular firm has a constant return to scale imposes exact restrictions on parameters of the model. Using such a model Kumar and Dube (2014) have demonstrated superiority of the restricted least squares estimator over the estimator that does not utilize prior information.

In an interesting article Mittlehammer (1984) has shown the admissibility of the restricted least squares estimator under predictive squared error risk. Use of Restricted Least Squares Estimator, in several other contexts, has also been reported by many authors over the years [see e.g. Dube and Singh (1995), Dube and Manocha (2002), Li ad Yang (2010) and Dube et. al (2015) among others and references cited therein].

Thus, it is generally perceived that an estimator which utilizes the prior information in account, a restricted estimator, will be more efficient than one which does not, if the prior information embodied in the imposed restrictions are correct. However, if the prior information is not correct the restricted estimator is likely to be more biased and may have a higher risk than unrestricted estimator. In situations when one is uncertain of the validity of the prior information and has some doubt as to whether or not to impose the restrictions on the model's parameter, a common procedure is to pre-test a priori the available information. Compatibility of sample information and prior information, for long, has been tested using Pre-test procedures [see; e.g Judge and Bock (1978), Ullah and Ullah (1981), and references cited therein]. Credit in this direction may be attributed to Bancroft (1944) who was first to use non-sample prior information in addition to the sample information. He examined the impact of pre-test of significance on subsequent estimation and inference in linear regression model. Working in similar scenario, Toro-Vizcarrondo and Wallace (1968) and later Wallace (1972) suggested practical procedure for determining the estimator to use based on a test of compatibility of sample and exact prior information in a regression model and in doing so, implied a pre-test estimator. Several authors have examined the sampling properties of the parameters of the linear regression model after a pre-test for linear restrictions on the coefficients [see e.g. Clark, Giles and Wallace (1987 a, b), Wan(1993), Hoque (2004) among others]. Their approach involves computing the bias and mean squared error of the Pre-test estimator.

Considering linear regression model with two explanatory variables, Giles and Srivastava (1993) obtained the exact distribution of least squares regression coefficient estimator after a preliminary t-test on one of the coefficients. Working with their idea, the aim of this paper is to determine the exact distribution of the Pre-test estimator assuming errors in the linear regression model to be normal. For this purpose orthonormal regression model is considered and the probability density function of an individual coefficient in the parameter vector is obtained. With the help of probability density function, moments of the Pre-test estimators of the coefficients are obtained which eventually lead to the computation of bias and risk of Pre-test estimator.

The plan of this paper is as follows: Section 2 describes the model and the estimators and utilization of tests of hypotheses for construction of Pre-test estimator of regression coefficients. In Section 3 the distribution function of the Pre-test estimator has been obtained with the help of which its density function is obtained and studied empirically. With the help of probability density function, the moments of the Pre-test estimators of the coefficients are obtained in Section 4. Lastly, in Appendix, proofs of theorems are provided.

2. THE MODEL AND PRE-TEST ESTIMATOR OF REGRESSION COEFFICIENTS

Considering the orthonormal linear regression model

$$y = Z\emptyset + \varepsilon \tag{2.1}$$

where y is an $n \times 1$ vector of observations on the response variable, Z is an $n \times p$ full column rank nonstochastic matrix of n observations on p explanatory variables, $\emptyset$ is a $p \times 1$ vector of unknown parameters associated with the p regressors and ε is an $n \times 1$ vector of disturbances, the elements of which are assumed to be independently and identically distributed each following normal distribution with mean zero and variance σ^2, so that $E(\varepsilon) = 0$, $E(\varepsilon\, \varepsilon') = \sigma^2 I_n$.

Following Judge and Bock (1978), the model (2.1) may be re-parameterized and we can convert the model to

$$y = X\beta + \varepsilon \tag{2.2}$$

where $X = Z(Z'Z)^{-1/2}$ and $\beta = (Z'Z)^{1/2}\,\emptyset$ so that $X'X = I_p$. Under this situation, the estimator of $\emptyset$ is $\widehat{\emptyset} = (Z'Z)^{1/2}\,\hat{\beta}$.

Under orthonormal model the least squares estimator from model (2.2) becomes $b = X'y$, which is normally distributed with mean $E(b) = \beta$ and variance $V(b) = \sigma^2 I_p$. The risk of b under quadratic error loss is given by

$$R(b) = E[(b-\beta)'(b-\beta)] = \sigma^2\, p \tag{2.3}$$

In order to construct a Pre-test estimator of regression coefficient in the presence of some prior information in the form of restrictions on the regression coefficients given by

$$r - R\beta = \delta \tag{2.4}$$

where r is a $q \times 1$ vector and R is $q \times p$ full row rank matrix of known elements, q being the number of restrictions imposed on the coefficients and δ is a $q \times 1$ vector representing the specification errors in the restrictions. In the orthonormal model, the application of least squares to (2.2) subjected to (2.4) leads to the following restricted least squares estimator of β given by

$$b_R = b + R'[RR']^{-1}(r - Rb) \qquad (2.5)$$

It is clear that owing to the normality of the estimators b the estimator b_R in (2.5) also follows Normal distribution with mean $E(b_R) = \beta + A\delta$; $A = R'(RR')^{-1}$ and with a variance covariance matrix $V(b_R) = \sigma^2[\mathrm{I} - R'(RR')^{-1}\mathrm{R}]$. The mean squared error of this estimator is $M(b_R) = \sigma^2[\mathrm{I} - R' RR'-1\mathrm{R} + A\delta\delta'A'$ so that the risk under quadratic error loss is given by

$$R(b_R) = \sigma^2(p - q) + 2\sigma^2\lambda \qquad (2.6)$$

where $\lambda = \frac{\delta'[RR']^{-1}\delta}{2\sigma^2}$ is non-centrality parameter. It is a measure of the validity of linear restrictions as it is monotonically related to the sum of squared errors in the individual restriction. Using this criterion to appraise performance, the restricted least squares estimator performs better than the ordinary least squares estimator if the difference between the risk of b and that of b_R is positive, i.e. $R(b) - \mathrm{R}(b_R) = \sigma^2(q - 2\lambda)$.

Hence, in this case, the restricted least squares estimator has smaller risk than the ordinary least squares estimator so long as $\lambda \leq \frac{q}{2}$. Since λ is always positive, the condition of dominance of b_R over b becomes $0 \leq \lambda \leq \frac{q}{2}$.

As, whether or not to incorporate the prior information may be decided by carrying out the following statistical test of significance:

$$H_0 : r - R\beta = \delta = 0, \qquad (2.7)$$

against the alternative

$$H_1 : r - R\beta = \delta \neq 0 \qquad (2.8)$$

The above hypothesis is conventionally tested using likelihood ratio procedure and leads to using the test statistic

$$u = \frac{v}{q} \frac{(r-Rb)'[RR']^{-1}(r-Rb)}{e'e} \qquad (2.9)$$

Under the null hypothesis (2.7), the test statistic u has a central F distribution with q and $v = n - p$ degrees of freedom. Alternatively under H_1, if the linear restrictions are not true, u has a non-central F distribution with q and $v = n - p$ degrees of freedom with non-centrality parameter λ. Following the likelihood ratio principle, the null hypothesis (2.7) is rejected, if the statistic u is greater than some critical value c, where the value of c is determined on the basis of the size α of the test and is determined from

$$\int_c^\infty dF(u) = \alpha$$

In view of this, a Pre-test estimator of regression coefficient may be defined as follows:

$$b_{PT} = I_{[0,c)}(u)b_R + I_{(c,\infty]}(u)b \qquad (2.10)$$

where c is the critical value of the pre-test and $I(.)$ is an indicator function which is one if u falls in the given interval and zero otherwise. Thus

$$I_{[0,c)}(u) = \begin{cases} 1 & if \quad 0 \le u < c \\ 0 & \text{otherwis} \end{cases}$$

It is easy to notice that $I_{(c,\infty]}(u) = 1 - I_{[0,c)}(u)$, so that the estimator (2.10) may be written as

$$b_{PT} = b - I_{[0,c)}(u)(b - b_R) \qquad (2.11)$$

3. DISTRIBUTION OF PRE-TEST ESTIMATOR OF REGRESSION COEFFICIENT

In order to study the performance properties of Pre-test estimator b_{PT}, given in (2.11), the statistic u can be written as the ratio of two Chi square distributed variables as follows:

$$u = \frac{u_1/q}{u_2/v} \;; \qquad v = n - p \qquad (3.1)$$

where $u_1 = \frac{(r-Rb)'\left[R\,R'\right]^{-1}(r-Rb)}{\sigma^2}$ and $u_2 = \frac{e'e}{\sigma^2}$ (3.2)

It is clear that under H_0, u_1 follows a Central Chi-square (χ^2) distribution with q degrees of freedom while under H_1 it follows Non-central Chi-square distribution with non-centrality parameter λ. It is also well known that u_2 is distributed independently of u_1 and follows a central χ^2distribution with v degrees of freedom [see; Searle (1971)].

Let us now first attempt to compute the distribution function and density function of the Pre-test estimator. Incidentally, derivation of the cumulative distribution function of b_{PT} also requires to know the distribution of $(b - b_R)$. As both b and b_R, follow Normal distribution, the difference $(b - b_R)$ also follows Normal distribution.

Hence $(b - b_R) \sim N[\delta^*, \Sigma]$ where $\delta^* = A\delta \; and \; \Sigma = (\sigma^2 AR)$. Now using these it is easy to find the distribution of the kth element of $(b - b_R)$. Thus it is observe that $\left(b_{R_k} - b_k\right) \sim N(\delta_k^*, {\sigma_*}^2)$ where δ_k^* is the kth element of the matrix Aδ and ${\sigma_*}^2$ is the kth diagonal element of the matrix $\sigma^2 R(R'R)^{-1}R'$. Having equipped with this information, to find the distribution function and density function of the Pre-test estimator first write the k^{th} element of the $p \times 1$ vector b_{PT} as

$$(b_{PT})_k = b_k - I\left(\frac{u_1}{u_2} < c^*\right)\left(b_k - b_{R_k}\right) \qquad (3.3)$$

where $c^* = \frac{q\,c}{v}$, b_k and b_{R_k} $(k = 1, 2, 3, \ldots\ldots, p)$ are the k^{th} elements of vectors b and b_R respectively.

It is worth mentioning here that we require the full sampling distribution of estimators not only to determine the size of critical region but also to study the implications of Pre-testing for interval estimation. Determination of the distribution is also pertinent for comparing different estimators on the basis of concentration probabilities. In view of this we have first obtained the distribution function of the Pre-test estimator (3.3) in the following theorem:

Theorem 1: When errors in the model (2.2) are normally distributed, the distribution function of the Pre-test estimator (3.3) is given by

$$F(\tau) = \Phi(\tau_1) - \sum_{i=0}^{\infty} w_i(\lambda) \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i, \frac{v}{2}\right) \Phi(\tau_2) \quad (3.4)$$

where $\tau_1 = \frac{\tau-\beta_k}{\sigma}$, $\tau_2 = \frac{\tau-\delta_k^*}{\sigma_*}$ and $\Phi(.)$ is the cumulative distribution function of standard normal distribution. The value $\mathrm{I}_x(a,b)$ is the incomplete Beta function and $w_i(\lambda) = \frac{e^{-\lambda}\lambda^i}{i!}$ are the Poisson weights given to the incomplete Beta function.

Proof: See Appendix.

From the above theorem, the density function of the Pre-test estimator may be obtained by differentiating $F(\tau)$ with respect to τ by applying Fundamental theorem of calculus. Hence, the probability density function of the k^{th} element of Pre-test estimator is given by

$$f(\tau) = \frac{1}{\sqrt{2\,\pi}}\left[\frac{1}{\sigma} e^{-\frac{1}{2}\left(\frac{\tau-\beta_k}{\sigma}\right)^2} - \varphi * \frac{1}{\sigma_*} e^{-\frac{1}{2}\left(\frac{\tau-\delta_k^*}{\sigma_*}\right)^2}\right] \quad ; \varphi = \sum_{i=0}^{\infty} w_i(\lambda)\ \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i, \frac{v}{2}\right) \quad (3.5)$$

The general interpretation of the above density function is obviously not straight forward. Therefore, for simplicity, the probability density function has been computed for some specific values of the parameters.

The Table 1 provides the probability density function for selected values of the parameters. These values are then graphed in Figure 1 for different values of λ in order to get an idea of the effect of the change in the non centrality parameter on the shape of the distribution of $(b_{PT})_k$. Analyzing the probability density function of k^{th} Pre-test estimator we notice from Figure 1 that the density function are almost identical for very small values of the non-centrality parameter. As the non-centrality parameter increases, the kurtosis of the distribution decreases while there is not much affect in the skewness of the distribution. It is also pertinent to mention that for $\tau \le -2$ and also for $\tau \ge 4$ the graph of the density function is almost coincident with X- axis.

Table 1 Density function of Pre-test estimator for selected values of $v = 30$, $q = 1$, $\beta_k = 1, \delta_k^* = 0.5, \sigma^2 = 1, \sigma_* = 0.1$

τ	λ=.01	λ=.1	λ=.5	λ=1	λ=5	λ=10
-3.50	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.00	0.0003	0.0003	0.0003	0.0003	0.0002	0.0001
-2.50	0.0020	0.0019	0.0017	0.0015	0.0010	0.0009
-2.00	0.0087	0.0085	0.0078	0.0071	0.0048	0.0045
-1.50	0.0307	0.0301	0.0279	0.0257	0.0187	0.0176
-1.00	0.0856	0.0842	0.0789	0.0736	0.0568	0.0542
-0.10	0.2991	0.2956	0.2820	0.2682	0.2251	0.2185
0.10	0.3558	0.3520	0.3369	0.3217	0.2741	0.2668
1.00	0.4847	0.4811	0.4667	0.4521	0.4066	0.3996
1.50	0.4110	0.4085	0.3986	0.3886	0.3574	0.3525
2.00	0.2735	0.2722	0.2669	0.2615	0.2448	0.2422
2.50	0.1427	0.1421	0.1399	0.1377	0.1307	0.1296
3.00	0.0583	0.0581	0.0574	0.0566	0.0544	0.0540
3.50	0.0186	0.0186	0.0184	0.0182	0.0176	0.0175
4.00	0.0046	0.0046	0.0046	0.0046	0.0045	0.0044

5.00	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
6.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Figure 1

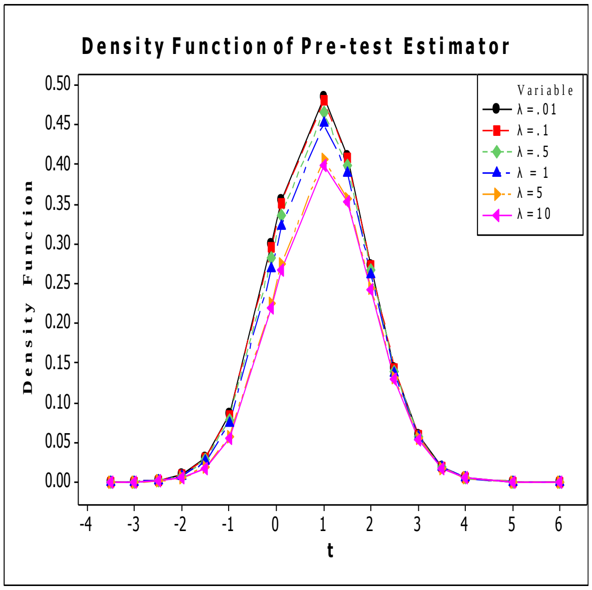
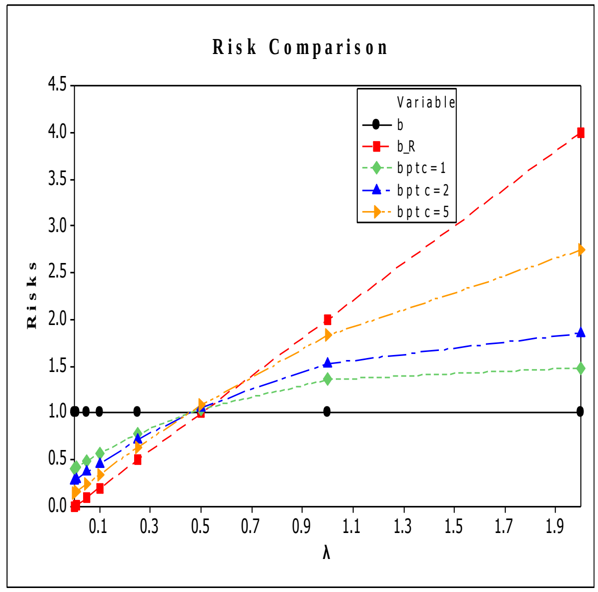


Figure 1 (A)



4. PROPERTIES OF PRE-TEST ESTIMATOR OF REGRESSION COEFFICIENT

The determination of distribution of the Pre-test estimator allows to compute moments of different order which will eventually lead to determine the bias, mean squared error and risk of Pre-test estimator $(b_{PT})_k$ under quadratic error loss. Thus using (3.5), the following theorem is obtained which gives the j^{th} Raw Moment of the Pre-test estimator.

Theorem 2 When errors in the model (2.2) are normally distributed, the j^{th} raw moment of the Pre-test estimator (3.3) is given by

$$E[(b_{PT})_k]^j = \frac{1}{\sqrt{2\,\pi}} \{\int_{-\infty}^{\infty} (\beta_k - \sigma g_1)^j e^{-\frac{g_1^2}{2}} dg_1$$

$$+ \sum_{i=0}^{\infty} w_i(\lambda) \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i, \frac{v}{2}\right) \int_{-\infty}^{\infty} (\delta_k^* - \sigma_* g_2)^j e^{-\frac{g_2^2}{2}} dg_2\} \qquad (4.1)$$

Proof: See Appendix.

With the help of the above Theorem, we can easily determine the bias and the risk under quadratic error loss function of the k^{th} element of the Pre-test estimator as follows:

$$B(b_{PT})_k = E[(b_{PT})_k - \beta_k] \qquad (4.2)$$

$$R(b_{PT})_k = E[(b_{PT})_k - \beta_k]^2 = E[(b_{PT})_k]^2 - 2\beta_k E[(b_{PT})_k] + {\beta_k}^2 \qquad (4.3)$$

Using (4.2) and (4.3), we get the following theorem which provides the bias and the Risk of the Pre test estimator $(b_{PT})_k$ under quadratic error loss.

Theorem 3: When the errors in the model (2.2) are normally distributed, the bias and the risk of the Pre-test estimator (3.3) under quadratic error loss is given by

$$B(b_{PT})_k = \delta_k^* \sum_{i=0}^{\infty} w_i(\lambda) \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i, \frac{v}{2}\right) \quad (4.5)$$

$$R(b_{PT})_k = \sigma^2 - ({\sigma_*}^2 + 2\,\sigma^2\lambda - 2\beta_k \delta_k^*) \sum_{i=0}^{\infty} w_i(\lambda) \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i, \frac{v}{2}\right) \quad (4.6)$$

Proof: See Appendix.

From (4.5) it is observed that when the k^th restriction is correct i.e. $\delta_k = 0$, the Pre-test estimator is unbiased. However, if restriction is not correct then bias of Pre-test is function of restriction specification error δ_k^*, the non centrality parameter λ, and the critical value c. It is also observed that the term $\sum_{i=0}^{\infty} w_i(\lambda) \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i, \frac{v}{2}\right) \leq 1$, so that the magnitude of the bias of the pre-test estimator is always smaller than that of the restricted least squares estimator.

Comparing the risks of k^th element of ordinary least squares and the Pre-test estimators, we see that the difference in the risks of the estimators is

$$\mathrm{R}(\mathrm{b}_{\mathrm{PT}})_{\mathrm{k}} - \mathrm{R}(\mathrm{b}_{\mathrm{k}}) = -({\sigma_*}^2 + 2\,\sigma^2\lambda - 2\beta_{\mathrm{k}}\delta_{\mathrm{k}}^*) \sum_{i=0}^{\infty} w_i(\lambda) \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i, \frac{v}{2}\right) \quad (4.7)$$

Hence, the Pre-test estimator and the least squares estimator have same risks if $\lambda = \frac{2\beta_k \delta_k^* - {\sigma_*}^2}{2\,\sigma^2}$. On the other hand, the least squares estimator has smaller risk than the Pre-test estimator when $\lambda > \frac{2\beta_k \delta_k^* - {\sigma_*}^2}{2\,\sigma^2}$. The Pre-test estimator has smaller risk than the least squares estimator so long as $< \frac{2\beta_k \delta_k^* - {\sigma_*}^2}{2\,\sigma^2}$. Interestingly, from the Theorem 3, it is easy to conclude that for c=0, the equation (4.6) provides the risk of k^th element of the ordinary least squares estimator as the value of Incomplete beta function becomes zero, while for large values of c it provides the risk of restricted least squares estimator. It is also worth mentioning here is that the risk of the ordinary least squares estimator remains constant and equals σ^2 as the estimator is unbiased. Moreover, it does not depend on the prior information hence, there is no effect of λ on the ordinary least squares estimator and the risk of ordinary least squares estimator remains constant for various values of λ. However, the risk of restricted least squares estimator changes as the value of λ changes while the risk of Pre-test estimator changes not only with the values of λ but also with changes in c, the critical value for a given level of significance. Therefore, in the empirical study firstly the risks of restricted least squares estimator and Pre-test estimator have been computed for different values of λ and c.

The risk differences have been graphed in Figure 1 (A) for different values of v, λ *and* c and are obtained for given values of parameters**.** The parameter values for numerical computations for comparing risks are $v = 20, q = 1,\ \beta_k = 1, \delta_k^* = 0.5, \sigma^2 = 1, \sigma_* = 0.1$. It is easy to see from Figure 1 (A) that the pre-test estimator is superior to the estimator based on sample information only over a small portion of the parameter space. When the prior information is true, the risk of the Pre-test estimator is larger than the risk of restricted least squares estimator. It remains so when the value of the non-centrality parameter λ lies between 0 to 0.5. This empirical result is justified by the result that the restricted least squares estimator has smaller risk than the ordinary least squares estimator so long as $\lambda \leq \frac{q}{2}$, also agreeing with the results obtained by Judge and Bock (1978). Therefore, the restricted least squares estimator dominates the ordinary least squares estimator if the restriction is correct, and the opposite is true when the restriction specification error increases. Now let us look at the effect of the change in the size of the critical region which is inversely related to the value of c. It is observed that the risk of the Pre-test estimator tends to the risk of ordinary least squares estimator when the value of c decrease, and conversely tends to the risk of the restricted least squares estimator when c is large. This in fact is the theoretical basis of the formulation of the pre-test estimator.

In order to understand the behaviour of risk properties of the pre-test estimator with the change in the amount of non-centrality parameter and the size of the critical region, the risks values have been tabulated for various values of v, λ *and* c. The Tables 2 to 5 present the results on risk performances of the Pre-test estimator.

From the Tables 2 and Table 3, it is easy to see that when the value of the non-centrality parameter is less than 0.25 and $v = 10$, the risk of the Pre-test estimator decreases monotonically as the values of c increase, however, an increasing trend is observed when the values of non- centrality parameter is larger than 0.25. Firstly the graphs have been plotted for all the tabulated values of c and λ which are shown in Figure 2 which presents a comprehensive view of all possible values of λ. As the graph does not provide a clear picture of the patterns obtained, the graphs have again been plotted for smaller and then for few selected values of λ. Figures 2 (A) shows some specific smaller values of λ. It is evident from the Figure 2 (A) that for small values of the non-centrality parameter, in the range $0 < \lambda < 0.25$, the risk of the Pre-test estimator decreases. From Figure 2 (B) it clear that the magnitude of risk difference increases with increase in λ.

Coming to the values computed in Table 3, it can be seen that an increase in the value of $\sigma_{*,}$ makes an increase in the risk of pre-test estimator. These values are graphically represented in Figures 3, 3 (A) and 3 (B).

Lastly let us study the effect of change in the degrees of freedom on the performance properties of pre-test estimator. In the Tables 4 and 5, the value of the degrees of freedom is changed from 10 to 50. It is important to notice that the behaviour of risk of the Pre-test estimator does not change if we increase degrees of freedom i.e *v*, this can be verified from Tables 4 and 5.

Table 2 Risk of Pre-test estimator for selected values of $\boldsymbol{\sigma^2 = 1, v = 10, q = 1, \beta_k = 1, \delta_k^* = .5}$ *and* $\boldsymbol{\sigma_* = .1}$

c	λ=0	λ=.05	λ=.1	λ=.25	λ=.5	λ=1	λ=5	λ=10
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.1	0.7825	0.8158	0.8464	0.9240	1.0149	1.1014	1.0177	1.0003
1	0.4068	0.4908	0.5698	0.7788	1.0463	1.3567	1.1491	1.0062
2	0.2689	0.3661	0.4591	0.7135	1.0630	1.5308	1.4055	1.0309
3	0.2025	0.3038	0.4019	0.6767	1.0734	1.6589	1.7706	1.0919
4	0.1660	0.2685	0.3685	0.6539	1.0804	1.7552	2.2191	1.2072
5	0.1444	0.2470	0.3478	0.6389	1.0853	1.8283	2.7219	1.3913
6	0.1309	0.2332	0.3342	0.6287	1.0888	1.8842	3.2521	1.6529
7	0.1220	0.2241	0.3251	0.6216	1.0914	1.9272	3.7880	1.9943
8	0.1161	0.2178	0.3188	0.6165	1.0933	1.9605	4.3133	2.4123
9	0.1120	0.2135	0.3143	0.6128	1.0947	1.9866	4.8167	2.8993
10	0.1091	0.2103	0.3110	0.6100	1.0957	2.0071	5.2912	3.4451
15	0.1028	0.2033	0.3036	0.6035	1.0984	2.0624	7.1593	6.6555

Table 3 Risk of Pre-test estimator for selected values of $\boldsymbol{\sigma^2 = 1, v = 10, q = 1, \beta_k = 1, \delta_k^* = .5}$ *and* $\boldsymbol{\sigma_* = .5}$

c	$\lambda=0$	$\lambda=.05$	$\lambda=.1$	$\lambda=.25$	$\lambda=.5$	$\lambda=1$	$\lambda=5$	$\lambda=10$
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	0.6704	0.7454	0.8156	1.0000	1.2317	1.4864	1.1557	1.0063
2	0.5938	0.6831	0.7682	1.0000	1.3148	1.7238	1.4233	1.0316
3	0.5570	0.6519	0.7437	1.0000	1.3670	1.8985	1.8045	1.0938
4	0.5367	0.6343	0.7294	1.0000	1.4022	2.0298	2.2727	1.2115
5	0.5247	0.6235	0.7205	1.0000	1.4267	2.1295	2.7976	1.3995
6	0.5171	0.6166	0.7147	1.0000	1.4441	2.2057	3.3511	1.6666
7	0.5122	0.6120	0.7108	1.0000	1.4569	2.2643	3.9106	2.0151
8	0.5090	0.6089	0.7081	1.0000	1.4663	2.3098	4.4589	2.4418
9	0.5067	0.6067	0.7061	1.0000	1.4733	2.3453	4.9845	2.9390
10	0.5051	0.6052	0.7047	1.0000	1.4787	2.3733	5.4798	3.4963
15	0.5015	0.6016	0.7015	1.0000	1.4921	2.4487	7.4300	6.7740

Figure 2

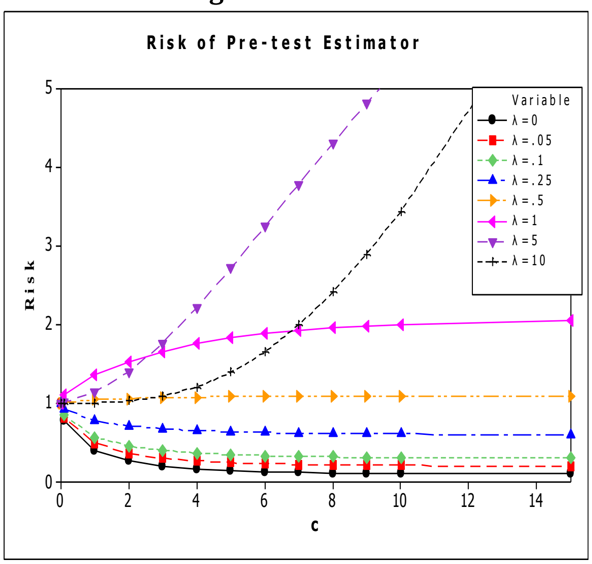


Figure 2 (A)

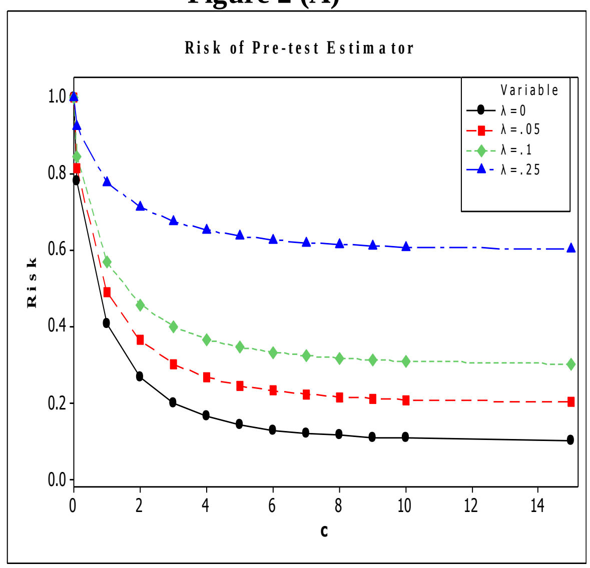


Figure 2 (B)

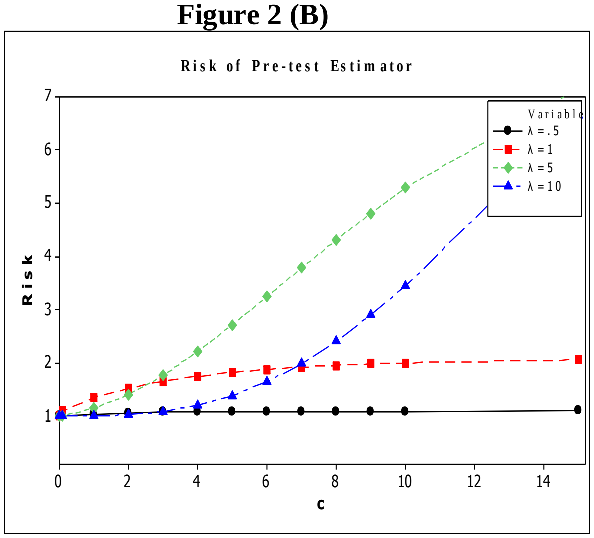


Figure 3

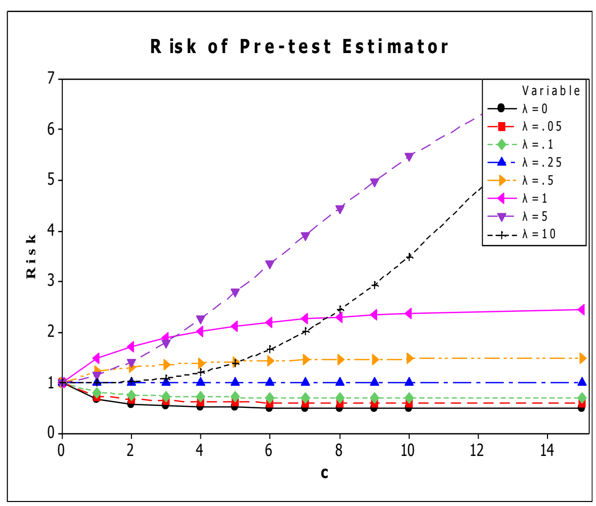


Figure 3 (A)

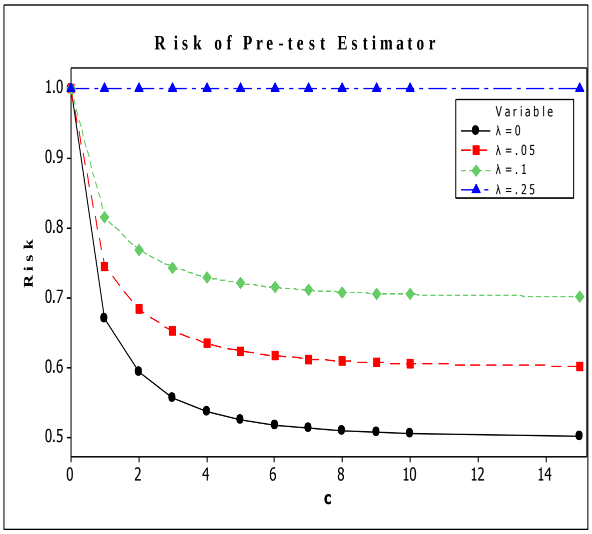


Figure 3 (B)

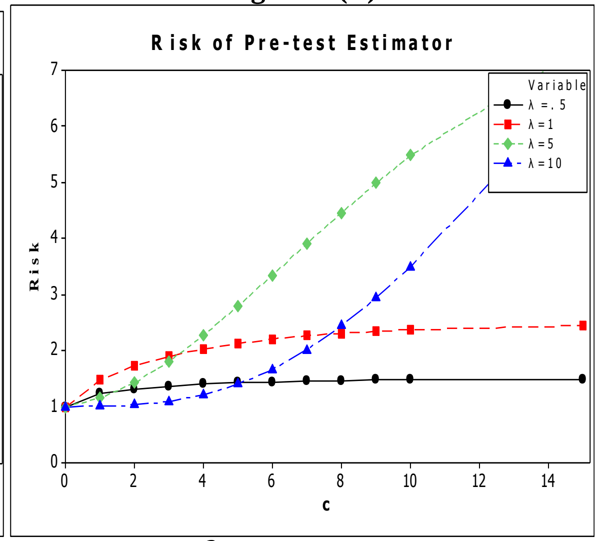


Table 4 Risk of Pre-test estimator for selected values of $\sigma^2 = 1, v = 50, q = 1, \beta_k = 1 \;\; \delta_k^* = .5 \; and \; \sigma_* = .1$

c	λ=0	λ=.05	λ=.1	λ=.25	λ=.5	λ=1	λ=5	λ=10
0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1	0.4139	0.482	0.5624	0.7752	1.047	1.3606	1.1441	1.0055
2	0.2756	0.3544	0.449	0.7082	1.0641	1.539	1.3862	1.0258
3	0.2096	0.2915	0.391	0.6705	1.0749	1.6722	1.7344	1.0742
4	0.1742	0.2567	0.3579	0.6473	1.0822	1.7732	2.1716	1.1659
5	0.154	0.2364	0.3379	0.6324	1.0872	1.85	2.6744	1.3149
6	0.1419	0.2239	0.3254	0.6226	1.0907	1.9082	3.2182	1.5327
7	0.1344	0.216	0.3174	0.616	1.0932	1.9524	3.7807	1.8266
8	0.1297	0.211	0.3121	0.6115	1.0949	1.986	4.3432	2.1994
9	0.1266	0.2076	0.3085	0.6083	1.0962	2.0116	4.891	2.6494
10	0.1246	0.2054	0.3061	0.6061	1.0972	2.0312	5.4137	3.1715
15	0.1209	0.2011	0.3013	0.6015	1.0992	2.079	7.4922	6.5248

Table 5 Risk of Pre-test estimator for $\sigma^2 = 1, v = 50, \; q = 1, \beta_k = 1, \; \delta_k^* = .5 \; and \; \sigma_* = .5$

c	λ=0	λ=.05	λ=.1	λ=.25	λ=.5	λ=1	λ=5	λ=10
0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1	0.6646	0.741	0.8125	1.000	1.2351	1.4918	1.1505	1.0056
2	0.5863	0.6772	0.7639	1.000	1.3205	1.735	1.4031	1.0263
3	0.5493	0.6457	0.739	1.000	1.3745	1.9166	1.7667	1.0758
4	0.5296	0.6284	0.7248	1.000	1.4108	2.0544	2.2231	1.1694
5	0.5185	0.6182	0.7162	1.000	1.4359	2.159	2.748	1.3215
6	0.5118	0.612	0.7109	1.000	1.4534	2.2384	3.3157	1.5439
7	0.5078	0.608	0.7074	1.000	1.4658	2.2987	3.9029	1.8439
8	0.5052	0.6055	0.7052	1.000	1.4747	2.3445	4.4901	2.2245
9	0.5035	0.6038	0.7036	1.000	1.4812	2.3794	5.0621	2.684
10	0.5024	0.6027	0.7026	1.000	1.4858	2.4061	5.6077	3.217
15	0.5005	0.6006	0.7006	1.000	1.4962	2.4713	7.7776	6.6405

Figure 4

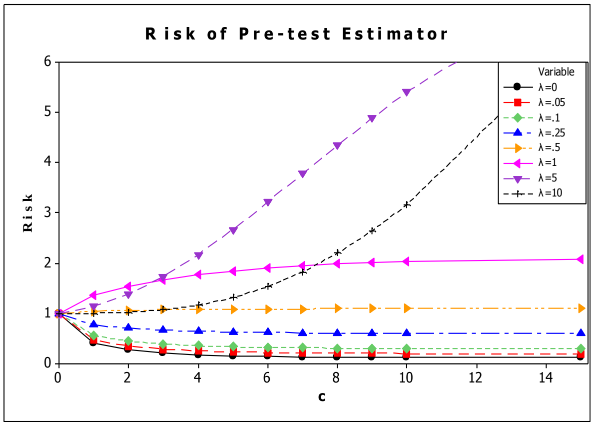
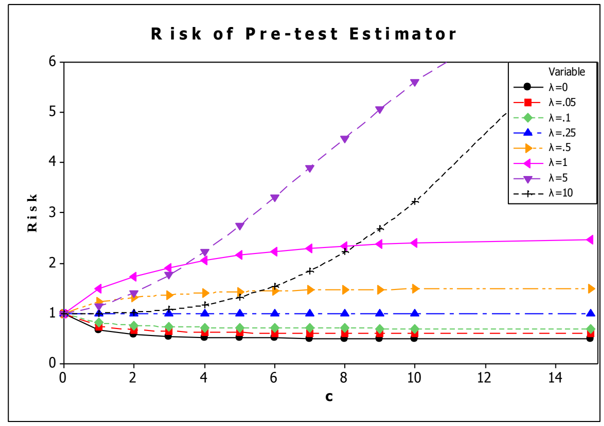


Figure 5



The values given in Tables 4 and 5 are presented in the Figures 4 and 5. It may be noted that although there is not a significant change in the risk pattern, the behaviour of risk is significantly different for $\lambda < 0.25$ and that for $\lambda > 5$, as seen in the case when the degrees of freedom were taken as 10. However it may be noted that the effect of σ_* is again visible as the slope of the curves sharply increase with the increase in its value.

Appendix

Proof of Theorem 1

As elements of ε in the model (2.2) are independently and identically distributed each with mean zero and variance σ^2 , considering the orthonormal version of the linear statistical model where $X'X = I_p$, the elements of b are independently and identically distributed each following normal distribution with mean β_k and variance σ^2. Using simple laws of probability, we can find the cumulative distribution function of $(b_{PT})_k$; $k = 1, 2, 3, \dots \dots \dots \dots p$ as follows. As

$$(b_{PT})_k = b_k - I\left(\frac{u_1}{u_2} < c^*\right)\left(b_k - b_{R\,k}\right)$$
$$F(\tau) = P[(b_{PT})_k \le \tau];$$
$$= P[b_k \le \tau] - P\left[\frac{u_1}{u_2} \le c^* \text{ and } z_k \le \tau\right]; \quad z_k = b_k - b_{R\,k} \qquad \text{(A.1)}$$

Owing to independence of $\frac{u_1}{u_2}$ and z_k , we can write (A.1) as

$$F(\tau) = P[b_k \le \tau] - P[\,z_k \le \tau].\,P\left[\frac{u_1}{u_2} \le c^*\right] = P_1 - P_2.P_3 \qquad \text{(A.2)}$$

It may be noted that P_1 and P_2 are nothing but the cumulative normal distribution functions and may be written as

$$P_1 = P[b_k \le \tau] = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\tau} e^{-\frac{1}{2}\left(\frac{b_k - \beta_k}{\sigma}\right)^2} db_k = \Phi(\tau_1) \; ; \quad \tau_1 = \frac{\tau - \beta_k}{\sigma} \; ; \qquad \text{(A.3)}$$

$$\text{and} \quad P_2 = P[z_k \le \tau] = \frac{1}{\sqrt{2\pi}\,\sigma_*}\int_{-\infty}^{\tau} e^{-\frac{1}{2}\left(\frac{\tau - \delta_k^*}{\sigma_*}\right)^2} dz_k \; P_2 = \Phi(\tau_2) \; ; \; \tau_2 = \frac{\tau - \delta_k^*}{\sigma_*} \qquad \text{(A.4)}$$

Lastly, the quantity P_3 involved in $F(\tau)$ is computed as follows:

$$P_3 = P\left[\frac{u_1}{u_2} \leq c^*\right] \quad ; \qquad c^* = \frac{qc}{v}$$

As u_1 follows Non-Central Chi-square distribution with non-centrality parameter λ and p degree of freedom while u_2 is central Chi-square distribution with $v = n - p$ degrees of freedom, from Judge and Bock (1978) [see; Theorem 1, Appendix B.3], we can write

$$P_3 = \sum_{i=0}^{\infty} w_i(\lambda)\ P\left[\frac{\chi^2_{q+2i}}{\chi^2_v} < c^*\right] \quad ; \quad w_i(\lambda) = \frac{e^{-\lambda}\lambda^i}{i!} \qquad \text{(A.5)}$$

Therefore, $P_3 = \sum_{i=0}^{\infty} w_i(\lambda)\ P\left[\frac{\chi^2_{q+2i}}{\chi^2_v} < c^*\right] = \sum_{i=0}^{\infty} w_i(\lambda)\ \frac{1}{B\left(\frac{q}{2}+i,\frac{v}{2}\right)} \int_0^{c^*} \frac{u^{\frac{q}{2}+i-1}}{(1+u)^{\frac{v+q}{2}+i}}\, du$ (A.6)

Using the transformation $\frac{u}{1+u} = t_1 \quad \Rightarrow \quad u = \frac{t_1}{1-t_1}$ (A.6) reduces to

$P_3 = \sum_{i=0}^{\infty} w_i(\lambda) \frac{1}{B\left(\frac{q}{2}+i,\frac{v}{2}\right)} \int_0^{c_1} t_1^{\ \frac{q}{2}+i-1}(1-t_1\ )^{\frac{v+q}{2}+i-1}\, dt_1$ where $c_1 = \frac{c^*}{1+c^*} = \frac{qc}{v+qc}$.

Hence, $P_3 = \sum_{i=0}^{\infty} w_i(\lambda)\, .\, \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i,\ \frac{v}{2}\right)$ (A.7)

where $\mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i,\ \frac{v}{2}\right) = \frac{1}{B\left(\frac{q}{2}+i,\frac{v}{2}\right)} \int_0^{\frac{qc}{v+qc}} t^{\frac{q}{2}+i-1}(1-t)^{\frac{v+q}{2}+i}\, dt$ is incomplete beta function. Substituting the values of P_1 , P_2 and P_3 from (A.3), (A.4) and (A.7) in (A.1) we get the result (3.4) in Theorem 1. Application of the Fundamental theorem of calculus to (3.4) leads to get the probability density function of $(b_{PT})_k$; $k = 1,2,3, \dots\dots\dots\dots p,$ the k^{th} element of the Pre-test estimator.

Proof of Theorem 2

Using the probability density function (3.5) the moments of order j of $(b_{PT})_k$ can be obtained. Noticing that

$$E[(b_{PT})_k]^j = \int_{-\infty}^{\infty} [(b_{PT})_k]^j\ f\ [(b_{PT})_k]\, d(b_{PT})_k$$

$$= \frac{1}{\sqrt{2\,\pi}} \left[\frac{1}{\sigma}\int_{-\infty}^{\infty} \tau^j\ e^{-\frac{1}{2}\left(\frac{\tau-\beta_k}{\sigma}\right)^2} d\tau - \sum_{i=0}^{\infty} w_i(\lambda)\ \mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i,\frac{v}{2}\right)\frac{1}{\sigma_*}\int_{-\infty}^{\infty} \tau^j\ e^{-\frac{1}{2}\left(\frac{\tau-\delta_k^*}{\sigma_*}\right)^2} d\tau\right] \text{(A.8)}$$

using the transformations $\frac{\tau-\beta_k}{\sigma} = g_1$ *and* $\frac{\tau-\delta_k^*}{\sigma_*} = g_2$ in (A.8) we can obtain the j^{th} order moment of $(b_{PT})_k$ as given in Theorem 2.

Proof of Theorem 3

The bias and the risk under quadratic error loss of the Pre-test estimator is given by

$$B(b_{PT})_k = E[(b_{PT})_k] - \beta_k \qquad \text{(A.9)}$$

$$R(b_{PT})_k = E[(b_{PT})_k]^2 - 2\beta_k E[(b_{PT})_k] + {\beta_k}^2 \qquad \text{(A.10)}$$

In order to find the bias and the risk under quadratic error loss of the Pre-test estimator let us substitute $j = 1$ and $j = 2$ in (A.8) as

$$E[(b_{PT})_k] = \frac{1}{\sqrt{2\,\pi}} \{\int_{-\infty}^{\infty}(\beta_k + \sigma g_1)e^{-\frac{g_1^2}{2}}dg_1$$

$$- \sum_{i=0}^{\infty} w_i(\lambda)\mathrm{I}_{\frac{qc}{v+qc}}\left(\frac{q}{2}+i,\frac{v}{2}\right) \int_{-\infty}^{\infty}(\delta_k^* + \sigma_* g_2)e^{-\frac{g_2^2}{2}}dg_2\} \qquad \text{(A.11)}$$

$$E[(b_{PT})_k]^2 = \frac{1}{\sqrt{2\,\pi}} \{\int_{-\infty}^{\infty}(\beta_k + \sigma g_1)^2 e^{-\frac{g_1^2}{2}}dg_1$$

$$- \sum_{i=0}^{\infty} w_i(\lambda)\mathrm{I}_{\frac{pc}{v+pc}}\left(\frac{p}{2}+i,\frac{v}{2}\right) \int_{-\infty}^{\infty}(\delta_k^* + \sigma_* g_2)^2 e^{-\frac{g_2^2}{2}}dg_2\} \qquad \text{(A.12)}$$

On solving the integrals in (A.11) and (A.12) we can obtain the bias and the risk under quadratic error loss function of the Pre-test estimator $B(b_{PT})_k$ as given in Theorem 3.

References

1. Bancroft, T. A. (1944), "On Biases in Estimation due to the use of Preliminary Tests of Significance", *Annals of Mathematical Statistics*, 15, 190-204.
2. Clarke, J.A., Giles D.E.A. and Wallace T.D. (1987a), "Estimating the Error Variance in Regression after a Preliminary test of Restrictions on the Coefficients", *Journal of Econometrics* 34, 293-304.
3. Clarke, J.A., Giles D.E.A. and Wallace (1987b), "Preliminary Test Estimation of the Error Variance in Linear Regression", *Econometric Theory* 3, 299-304.
4. Dube, M. and Manocha, V. (2002): “Simultaneous Prediction in Restricted Regression Models.” *Journal of Applied Statistical Science*, 11(4), 277-288.
5. Dube, M. and Singh V. (1995), “Predictive Efficiency of Improved Estimators in Restricted Regression Models” *Test, 4(2), 323-331.*
6. Dube, M., Kumar V. and Isha (2015), “The Performance Properties of Disturbance Variance under Restrictions using LNEX Loss Function” *Journal of Applied Probability and Statistics*, 10(02), pp. 79-88
7. Giles, D. E. A. and Srivastava V. K. (1993), "The exact Distribution of a Least Squares Regression Coefficient Estimator after a Preliminary-t Test", *Statistics and Probability Letters* 16, 59-64.
8. Hoque, Z. (2004), “Improved Estimation of Linear Models under Different Loss Functions, Ph. D. Thesis, University of Southern Queensland, Australia
9. Judge, G. G. and Bock, M. E. (1978). The *Statistical Implications of Pre-test andstein-Rule Estimators in Econometrics*, North Holland, New York.
10. Kumar V. and Dube M (2014), “Use of Restricted Least Squares Estimator to Study the Growth of Manufactured Food Products in Haryana”, *International Journal of Statistika and Mathematika*, 9(1), 2014, 48-49
11. Li, Y. and Yang, H. (2010), “Two Kinds of Restricted Estimators in Singular Linear Regression”, *Communication in Statistics- Theory and Methods*, 39, 2594-2603.
12. Mittlehammer, R. C. (1984), “On the Admissibility of restricted Least Square Estimator”, *Communication in Statistics, Theory and Methods*, 13, 1135-1146.
13. Searle, S. R. (1971). “Linear Models”. Wiley, New York
14. Toro-Vizcarrondo, C. E. and Wallace T. D. (1968), “A Test of the Mean Squared Error Criterion for Restrictions in Linear Regression”, *Journal of the American statistical Association*, 63, 558-572.

15. Vinod, H.D. and Ullah, A. (1981), "*Recent Advances in Regression Analysis*", Marcel Dekker, New York.
16. Wallace T. D. (1972), "Weaker Criteria and Tests for Linear Restrictions in Regression", *Econometrica, 40, 689-698.*
17. Wan A. T. K. (1993), "*Inequality Restricted and Pre-test Estimation an a Mis-Specified Econometric Model*" Ph. D. Thesis, University of Canterbury, Christchurch, New Zealand.